Sensitivity Analyses Comparing Time-to-Event Outcomes Only Existing in a Subset Selected Post-Randomization

Bryan E. Shepherd
Department of Biostatistics, Vanderbilt University, Nashville, TN, 37232, USA
bryan.shepherd@vanderbilt.edu

Peter B. Gilbert, Thomas Lumley
Department of Biostatistics, University of Washington, Seattle, WA, 98195, USA

In some randomized studies researchers are interested in determining the effect of treatment assignment on outcomes that may only exist in a subset chosen after randomization. For example, in preventative HIV vaccine efficacy trials it is of interest to determine whether randomization to vaccine affects post-infection outcomes that may be right censored. A few such outcomes in these trials include time from infection diagnosis to initiation of antiretroviral therapy or time from infection diagnosis to AIDS. Here we present sensitivity analysis methods for making causal comparisons on these post-infection outcomes. We focus on estimating the survival causal effect, defined as the difference between probabilities of not yet experiencing the event in the vaccine and placebo arms, conditional on being infected regardless of treatment assignment. This group is referred to as the always infected principal stratum. Our key assumption is monotonicity, that subjects randomized to the vaccine arm who become infected would have been infected if randomized to placebo. We propose nonparametric, semiparametric, and parametric methods for estimating the survival causal effect. These methods are applied to the first Phase III preventative HIV vaccine trial,
VaxGen’s trial of AIDSVAX B/B.

KEYWORDS: Causal inference; Principal stratification; Kaplan-Meier; AIDS.

1. INTRODUCTION

In preventative HIV vaccine efficacy trials it is of interest to assess vaccine effects on outcomes that occur only in infected individuals (Nabel 2001; Graham 2002; Gilbert et al., 2005). The current wave of candidate HIV vaccines undergoing efficacy testing have been specifically designed to alter disease progression and post-infection outcomes (www.hvtn.org). The effect of the vaccine on time-to-event post-infection outcomes, such as the time from HIV infection diagnosis until AIDS, are particularly of interest.

In efficacy trials, such as VaxGen’s trial of AIDSVAX B/B, thousands of healthy HIV uninfected volunteers are randomized to vaccine or placebo. They are followed for a certain amount of time, with the primary endpoint being HIV infection. Those who get infected are then enrolled into a post-infection study, and monitored for several additional years.

Comparisons of outcomes between those in the vaccine and placebo arms who become infected during the course of the trial could be misleading, because they condition on a post-randomization variable (infection) and therefore are susceptible to selection bias (Rosenbaum 1984, Halloran and Struchiner 1995). One could perform an intention-to-treat (ITT) analysis using all randomized individuals, ignoring infection status and defining the outcome of interest as the post-infection outcome (e.g., time from randomization until AIDS). However, this is problematic because the majority of trial participants will not get infected and are
therefore not followed past the time that the initial study concludes, $\tau_0$. If one censors these uninfected individuals at $\tau_0$, then one induces dependent censoring and the Kaplan-Meier method provides a biased estimate that under-estimates the survival probability for $t > \tau_0$. On the other hand, if one assumes that the post-infection outcome occurs in none of the individuals who are not infected at time $\tau_0$ and censors these individuals at the time when the post-infection study ends, $\tau_0 + \tau$, then one is most likely over-estimating the survival probability for $t > \tau_0$, particularly if $\tau$ is large. An unbiased ITT analysis would be to censor everyone who does not experience the post infection event at $\tau_0$. This is inefficient because much information is thrown away. From a statistical standpoint, the ideal trial design would not have an initial stopping point, but would follow all individuals for the entire study time, $\tau_0 + \tau$. However, resource constraints make an efficacy trial of this type infeasible.

The method we propose here is based on principal stratification (Frangakis and Rubin 2002). Our analysis addresses a different question than an ITT analysis: among those who would have been infected regardless of treatment assignment, does assignment to the vaccine increase/decrease the probability of being AIDS free $t$ months post-infection? It has been pointed out by many that comparing an outcome in the subgroup of individuals who would have been infected regardless of treatment assignment is a causal comparison (Kalbfleish and Prentice 1980; Robins 1995; Rubin 2000; Robins and Greenland 2000; Frangakis and Rubin 2002). This subgroup has been referred to as the *always infected* or *doomed* principal stratum (Hudgens, Hoering, and Self 2003; Gilbert, Bosch, and Hudgens 2003; Hudgens and Halloran 2006). An analysis conditioning on membership in the always infected principal stratum is
particularly of interest when one wants to know whether or not there exists a mechanism through which the vaccine alters a time-to-event outcome in infected individuals.

Because one does not known which participants would have been infected regardless of treatment assignment, assumptions must be made to answer the causal question. Our key assumption is that any participant in the vaccine arm who becomes infected would have been infected if randomized to placebo. Under this assumption, we advocate a sensitivity analysis approach where the probability of infection if assigned vaccine, given infection in the placebo arm and time from infection diagnosis to some event of interest, is of a known form indexed by an assumed sensitivity parameter.

This work extends the sensitivity analysis methods reported in Hudgens et al. (2003) (HHS); Gilbert et al. (2003) (GBH); and Shepherd, Gilbert, Jemiai, and Rotnitzky (2006) (SGJR) to a time-to-event outcome defined post-randomization. In Section 2 we discuss notation, assumptions, and estimands; in Section 3 we discuss estimation; in Section 4 we explore the finite sample properties of our estimators; and in Section 5 we apply our methods to investigate the effect of vaccination on the time to initiation of antiretroviral therapy in VaxGen’s trial of AIDSVAX B/B. Section 6 discusses the methods and results. Technical details are found in the Appendix.

2. NOTATION, CAUSAL ESTIMAND, AND ASSUMPTIONS

Consider a study in which $N$ subjects, independently and randomly selected from a given population of interest, are randomized to either placebo or vaccine. Let $Z_i = 1$ if
subject $i$, $i = 1, \ldots, N$, is randomized to vaccine and $Z_i = 0$ if randomized to placebo. Trial participants are monitored for HIV infection for a predetermined period of time; the infection status during the study follow-up period for subject $i$ is the indicator $S_i$ where $S_i = 1$ if infected and $S_i = 0$ if not. Let $T_i$ be the time from infection diagnosis until some event for subject $i$, and define $C_i$ as the time from infection diagnosis until censoring. If $S_i = 1$, we observe $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(Y_i = T_i)$. Notice that $T_i, C_i, Y_i$, and $\Delta_i$ only exist if $S_i = 1$; otherwise they are assigned the value $\ast$.

To define the estimand of interest, we use potential outcomes/counterfactuals (Neyman 1923, Rubin 1978, Robins 1986). Specifically, define $S_i(0)$ to be the infection status indicator if, possibly contrary to fact, subject $i$ is assigned placebo. Define $S_i(1)$ to be the infection status indicator if subject $i$ is assigned vaccine. Similarly, define $T_i(0)$ to be the time-to-event outcome if participant $i$ is assigned placebo and $T_i(1)$ the time-to-event outcome if assigned vaccine. The potential outcomes $C_i(z), Y_i(z)$, and $\Delta_i(z)$ are similarly defined for $z = 0, 1$. For a subject who does not become infected if assigned treatment $z$, i.e., $S_i(z) = 0$, we define $T_i(z) = Y_i(z) = C_i(z) = \Delta_i(z) = \ast$. This notation implicitly assumes that the potential outcomes of each trial participant are not influenced by the treatment assignments of other participants, an assumption known as the Stable Unit Treatment Value Assumption (SUTVA) (Rubin 1978). SUTVA may be violated in vaccine studies of infectious disease when a local population is treated intensively, producing herd immunity. It is a reasonable assumption when study participants are geographically dispersed, as in our example.

Assuming the study participants make up a random sample from a large population
of interest, the outcomes \( W_i = (Z_i, S_i(0), T_i(0), T_i(1), C_i(0), C_i(1)), \ i = 1, \ldots, N, \)
are i.i.d. copies of a random vector \( W = (Z, S(0), S(1), T(0), T(1), C(0), C(1)) \), and similarly the observed data \( O_i = (Z_i, S_i, Y_i, \Delta_i), \ i = 1, \ldots, N, \) are i.i.d. copies of \( O = (Z, S, Y, \Delta) \). Randomization ensures that

\[
(S(0), S(1), T(0), T(1), C(0), C(1)) \perp Z
\]

because \((S(0), S(1), T(0), T(1), C(0), C(1))\) can be considered an unobserved baseline characteristic of each subject. Here, for random variables \( A, B \) and \( C, A \perp B \mid C \) indicates conditional independence of \( A \) and \( B \) given \( C \).

The four basic principal strata (Frangakis and Rubin 2002) can be defined in terms of the counterfactual pair \((S(0), S(1))\): the never infected are those with \( S(0) = S(1) = 0 \), the harmed are those with \( S(0) = 0 \) and \( S(1) = 1 \), the protected are those with \( S(0) = 1 \) and \( S(1) = 0 \), and the always infected (a\( i \)) are those with \( S(0) = S(1) = 1 \).

For a subject \( i \) who is in the a\( i \) principal stratum, a causal effect on his/her time-to-event outcome is some measure of discrepancy between \( T_i(0) \) and \( T_i(1) \). For example, the difference \( T_i(1) - T_i(0) \) was used by HHS and GBH, with the average causal effect in the a\( i \) stratum defined as \( ACE = E[T(1) - T(0) \mid S(0) = S(1) = 1] \). With right-censored outcomes, such an estimand may not be identifiable without modeling the distribution of the time-to-event outcome; alternatively, one could compare restricted means (Chen and Tsiatis 2001). However, it is often of primary interest to compare the probabilities of being event-free at time \( t \) post-infection diagnosis, defined in the a\( i \) stratum as \( 1 - P(T_i(z) \leq t \mid S_i(0) = S_i(1) = 1) \) for \( z = 0, 1 \) (where \( t = 0 \) corresponds to the time of infection diagnosis). We take
this approach; our estimand of interest is the “survival” causal effect in the ai stratum:

\[ SCE(t) = P(T_i(0) \leq t | S_i(0) = S_i(1) = 1) - P(T_i(1) \leq t | S_i(0) = S_i(1) = 1) \]

\[ \equiv F_{ai}^p(t) - F_{ai}^v(t). \]

In order to estimate \( SCE(t) \), we will make the following key assumptions:

A.1: Monotonicity: \( S_i(1) \leq S_i(0) \) for all \( i \).

A.2: Independent Censoring: \( C_i(z) \perp T_i(z) | S_i(z) = 1 \) for \( z = 0, 1 \).

The first of these assumptions implies that no individual is in the harmed principal stratum. This is a strong but plausible assumption in randomized double-blinded placebo-controlled vaccine trials. The second assumption is common when analyzing time-to-event data, with the only difference being that this independence is made at the counterfactual level and is conditional on infection (otherwise \( C_i(z) = T_i(z) = * \)).

From these assumptions, \( F_{ai}^v(t) = F_v(t) \equiv P(T_i \leq t | Z_i = 1, S_i = 1) \), the distribution of the time-to-event outcome given vaccine and infection, and is therefore identified. However, in order to estimate \( SCE(t) \), we must also be able to identify \( F_{ai}^p(t) \) which is still not identifiable. The identifiable distribution, \( F_p(t) \equiv P(T_i \leq t | Z_i = 0, S_i = 1) \), is a mixture of the distribution of the time-to-event outcome for placebos in the always infected stratum, \( F_{ai}^p(t) \), and in the protected principal stratum, \( F_{prot}^p(t) \). The mixing parameter is one minus the probability of being in the always infected principal stratum given infection in the placebo arm, i.e., \( 1 - P(S(1) = 1 | S(0) = 1) \), which under A.1 equals \( 1 - P(S(1) = 1) / P(S(0) = 1) \equiv VE \), an oft-used measure of vaccine efficacy.
The distribution of $F_{ai}(t)$ can be represented as a biased sample from the distribution of the time-to-events for subjects with $S(0) = 1$:

$$F_{ai}(t) = (1 - VE)^{-1} \int_0^t w(s)dF_p(s),$$

where $w(t) \equiv P(S(1) = 1|S(0) = 1, T(0) = t)$. Bounds for $F_{ai}(t)$ are obtained by setting $w(t) = I_{t \leq q^{1-VE}}$ and $w(t) = I_{t \geq q^{VE}}$, with $q^{1-VE}$ and $q^{VE}$ being the $(1 - VE)^{th}$ and $VE^{th}$ quantiles of $T(0)$, respectively. This is equivalent to writing

$$F_{p,U}^{ai}(t) = \min \left\{ \frac{F_p(t)}{1 - VE}, 1 \right\},$$

$$F_{p,L}^{ai}(t) = \max \left\{ \frac{F_p(t) - VE}{1 - VE}, 0 \right\},$$

as expressions for the bounds of $F_{p}^{ai}(t)$, given by HHS.

There might be scientific reasons why these sharp bounds are too conservative. A different sensitivity analysis approach would use subject-matter knowledge to restrict the possible range for $F_{p}^{ai}(t)$ by choosing $w(t)$ resulting in $F_{p}^{ai}(t)$ somewhere between the bounds $F_{p,U}^{ai}(t)$ and $F_{p,L}^{ai}(t)$. This was the approach of GBH and is specified by making the following additional assumption:

A.3: $P(S(1) = 1|S(0) = 1, T(0)) = w(T(0); \beta)$, where $w(t; \beta) = \Phi \{\alpha + g(t; \beta)\}$, $\beta$ is fixed and known, $\Phi (\cdot)$ is a known cdf, $\alpha$ is an unknown parameter, and for each $\beta$, $g(t; \beta)$ is a known function of $t$.

If one assumes A.3, then $F_{p}^{ai}(t)$ is identified. Of course, the parameter $\beta$ is not identified by the observed data. It is regarded as fixed and known, and then varied over a plausible range of values as a form of sensitivity analysis (Scharfstein, Rotnitzky, and Robins 1999;
GBH). One choice for $w(t; \beta)$ is the expit function, i.e. $w(t; \beta) = (1 + \exp(-\alpha - \beta t))^{-1}$ used by GBH. For this choice of $w(\cdot)$, the sensitivity parameter $\beta$ has a log odds ratio interpretation. Using constant $w(\cdot)$ (e.g., the expit function with $\beta = 0$) is the same as assuming $T(0) \perp S(1)|S(0) = 1$, or equivalently that the distribution of the time-to-event outcome under placebo is the same in the ai and in the protected principal strata. Choosing $\beta = -\infty$ and $\beta = \infty$ corresponds to the sharp bounds given by (2) and (3).

3. ESTIMATION

We will first discuss nonparametric inference for the bounds of $SCE(t)$. We will then discuss parametric and semiparametric inference for $SCE(t)$ under assumed values of $\beta$.

3.1 Nonparametric Estimation: Sharp Bounds

Under SUTVA, (1), A.1, and A.2, and for $t < \tau$, where $\tau$ is the maximum post-randomization follow-up time, the sharp bounds given by (2) and (3) are consistently estimated as

$$
\hat{F}_{ai,U}^v(t) = \min \left\{ \frac{\hat{F}_p(t)}{1 - \overline{VE}}, 1 \right\},
$$

$$
\hat{F}_{ai,L}^v(t) = \max \left\{ \frac{\hat{F}_p(t) - \overline{VE}}{1 - \overline{VE}}, 0 \right\},
$$

where $\overline{VE} = \min\{1 - (n_vN_p)/(n_pN_v), 0\}$ ($n_z$ [n_z] is the number of individuals randomized to [infected in] treatment arm $z$, where $N_v + N_p = N$) and $\hat{F}_p(t)$ is the Kaplan-Meier estimator of $F_p(t)$. Since under monotonicity, $F_{ai}^v(t) = F_v(t)$, estimated sharp bounds for $SCE(t)$ are $\hat{F}_{ai,L}^v(t) - \hat{F}_v(t)$ and $\hat{F}_{ai,U}^v(t) - \hat{F}_v(t)$, where $\hat{F}_v(t)$ is the Kaplan-Meier estimator of $F_v(t)$. 

9
Under certain conditions, $\hat{F}^{ai,U}_p(t)$ and $\hat{F}^{ai,L}_p(t)$ are asymptotically normal. Specifically, $\hat{V}E$ is asymptotically normal if $0 < V E < 1$, or equivalently $0 < P(S(1) = 1) < P(S(0) = 1)$. The Kaplan-Meier estimate, $\hat{F}_p(t)$, is asymptotically normal under the usual conditions (Kalbfleish and Prentice 1980). Therefore, $\hat{F}^{ai,U}_p(t)$ and $\hat{F}^{ai,L}_p(t)$ are asymptotically normal if, in addition to these conditions, $0 < F_p(t) < 1 - V E$ and $V E < F_p(t) < 1$, respectively. (These latter conditions are a result of (2) and (3). If these conditions are violated then, by definition, $\hat{F}^{ai,U}_p(t)$ and $\hat{F}^{ai,L}_p(t)$ are 1 and 0, respectively, and hence estimates will not be asymptotically normal.) This asymptotic result follows from the asymptotic normality of $(\hat{F}_p(\cdot), \hat{V}E)$, the Hadamard differentiability of the maps $\hat{F}_p(t)/(1 - \hat{V}E)$ and $(\hat{F}_p(t) - \hat{V}E)/(1 - \hat{V}E)$, and the functional delta method (Andersen, Borgan, Gill, and Keiding, 1992). Under these conditions, expressions for the asymptotic variance are

\[
\text{var}\left(\hat{F}^{ai,U}_p(t)\right) = \left(\frac{p_0}{p_1}\right)^2 \sigma^2(t) + \left(\frac{F_p(t)}{p_1}\right)^2 \frac{p_0(1 - p_0)}{N_p} + \left(\frac{F_p(t)p_0}{p_1^2}\right)^2 \frac{p_1(1 - p_1)}{N_v},
\]

\[
\text{var}\left(\hat{F}^{ai,L}_p(t)\right) = \left(\frac{p_0}{p_1}\right)^2 \sigma^2(t) + \left(\frac{1 - F_p(t)}{p_1}\right)^2 \frac{p_0(1 - p_0)}{N_p} + \left(\frac{(1 - F_p(t))p_0}{p_1^2}\right)^2 \frac{p_1(1 - p_1)}{N_v},
\]

where $\sigma^2(t)$ is the variance of the Kaplan-Meier estimate $\hat{F}_p(t)$, $p_0 \equiv P(S = 1|Z = 0)$, and $p_1 \equiv p_0(1 - V E) = P(S = 1|Z = 1)$. From these equations one may estimate variances in the usual manner, by plugging in parameter estimates.

Variances may alternatively be estimated using a standard bootstrap procedure. Specifically, from $(O_1, \cdots, O_N)$ sample with replacement $N$ vectors $O_i$, creating $(O_1^*, \cdots, O_N^*)$. Compute $\hat{SCE}^*(t)$ based on the bootstrap sample $(O_1^*, \cdots, O_N^*)$. Repeat this process $K$ times, and estimate the variance of $\hat{SCE}(t)$ by the sample variance of the $\hat{SCE}^*(t)$s. Then 100(1 - $\alpha$)%-level confidence intervals may be constructed using the resulting variance esti-
mate (Wald-intervals), by the $\alpha/2$- and $(1-\alpha/2)$-quantiles of $\hat{SCE}(t)$ (percentile intervals), or by studentizing with the asymptotic variance estimate.

3.2 Parametric Estimation

In addition to SUTVA, (1), A.1, and A.2, assume A.3, and make the following modeling assumption:

- **M.1**: The distribution of $T(1)$ given $S(1) = 1$ is known up to a finite dimensional parameter $\eta_1$; that is, $f_{T(1)|S(1)=1}(t|S(1) = 1) = f_v(t; \eta_1)$ where $\eta_1$ is unknown and for each $\eta_1$, $f_v(\cdot; \eta_1)$ is a known density.

Also make one of the following two assumptions on the distribution of $T(0)$:

- **M.2a**: The distribution of $T(0)$ given $S(0) = 1$, is known up to a finite dimensional parameter $\eta^a_0$; that is, $f_{T(0)|S(0)=1}(t|S(0) = 1) = f_p(t; \eta^a_0)$, where $\eta^a_0$ is unknown and for each $\eta^a_0$, $f_p(\cdot; \eta^a_0)$ is a known density.

- **M.2b**: The distribution of $T(0)$ given $S(0) = S(1) = 1$ is known up to a finite dimensional parameter $\eta^b_0$; that is, $f_{T(0)|S(0)=S(1)=1}(t|S(0) = S(1) = 1) = f^{ai}_p(t; \eta^b_0)$, where $\eta^b_0$ is unknown and for each $\eta^b_0$, $f^{ai}_p(\cdot; \eta^b_0)$ is a known density.

For ease of reference, we call the model defined by SUTVA, (1), A.1, A.2, A.3, M.1, and M.2a model $\mathcal{M}_a$. We call $\mathcal{M}_b$ the model defined like $\mathcal{M}_a$ except replacing M.2a with M.2b and demanding that $w(t; \beta) > 0$ for all $t$.

Under $\mathcal{M}_a$, $SCE(t)$ is a function of unknown parameters $(\alpha, \eta^a_0, \eta_1)$. Specifically, $SCE(t) = \ldots$
\(SCE_a(t; \alpha, \eta_0^a, \eta_1)\) where

\[
SCE_a(t; \alpha, \eta_0^a, \eta_1) = \int_0^t w(s; \alpha, \beta) f_p(s; \eta_0^a) ds - \int_0^t f_v(s; \eta_1) ds.
\]

Similarly, under model \(M_b\), \(SCE(t) = SCE_b(t; \eta_0^b, \eta_1) = \int_0^t f^{ab}(s; \eta_0^b) ds - \int_0^t f_v(s; \eta_1) ds\). The maximum likelihood estimators of \(SCE(t)\) under \(M_a\) and \(M_b\) are therefore equal to the functions \(SCE_a(t; \cdot, \cdot, \cdot)\) and \(SCE_b(t; \cdot, \cdot, \cdot)\) evaluated at the ML estimators of \((\alpha, \eta_0^a, \eta_1)\) and \((\eta_0^b, \eta_1)\) under models \(M_a\) and \(M_b\), respectively.

In the absence of censoring, the likelihood induced by these assumptions (minus A.2) is given in SGJR. This likelihood can be easily modified to account for independent post-infection censoring:

\[
L_a(\alpha, \eta_0^a, \eta_1) \propto \prod_{i=1}^N \left\{ \left[ f_v(y_i; \eta_1) \right]^{s_i} (1 - F_v(y_i; \eta_1))^{1-s_i} p_0 \int w(y; \alpha, \beta) f_p(y; \eta_0^a) dy \right\}^{s_i} \times \left[ 1 - p_0 \int w(y; \alpha, \beta) f_p(y; \eta_0^a) dy \right]^{1-s_i} \times \left\{ f_p(y_i; \eta_0^a) \right\}^{s_i} (1 - F_p(y_i; \eta_0^a))^{1-s_i} p_0 \right\}^{1-s_i},
\]

under model \(M_a\). Under model \(M_b\) the likelihood \(L_b(\cdot)\) is defined as \(L_a(\cdot)\) but with \(f_p(y; \alpha, \eta_0^b) \equiv \frac{w^{-1}(y; \alpha, \beta) f^{ab}(y; \eta_0^b)}{\int w^{-1}(y; \alpha, \beta) f^{ab}(y; \eta_0^b) dy}\) replacing \(f_p(y; \eta_0^a)\).

Provided the protected principal stratum \(\{S_i(1) = 0, S_i(0) = 1\}\) is non-empty, then under sufficiently smooth parameterizations the ML estimators of the model parameters are asymptotically normally distributed. The variance of the normal limiting distribution can be consistently estimated with either the observed or the (estimated) expected information. These, in turn, can be used in conjunction with the delta method to obtain consistent
variance estimators of \( SCE(t) \) for each fixed \( t \). Sensitivity analyses are performed by varying \( \beta \).

It is worth noting that this likelihood can easily be extended to condition on baseline covariates. Likelihood based methods when the outcome of interest is not right censored, with or without baseline covariates, were extensively studied in SGJR; general principles stated there apply here.

### 3.3 Semiparametric Estimation

Consider estimation under SUTVA, (1), A.1, A.2, and A.3; i.e., performing sensitivity analyses by modeling \( w(\cdot) \) but leaving the distributions of \( T(0) \) and \( T(1) \) unspecified. One can think of this as extending GBH to time-to-event outcomes.

The estimating equations GBH used to estimate \( F^{ai}_p(t) \) when \( n_v/N_v < n_p/N_p \) can be written as

\[
0 = \psi(p_0, \alpha) = \begin{cases} 
\sum_{i=1}^{N}(1 - Z_i)(S_i - p_0) \\
\sum_{i=1}^{N} Z_i \left( S_i - p_0 \int_0^{\infty} w(t; \alpha, \beta) d\hat{F}_p(t) \right).
\end{cases}
\] (5)

When \( n_v/N_v \geq n_p/N_p \), \( F^{ai}_p(t) \) is estimated with \( \hat{F}_p(t) \), the nonparametric estimate of \( F_p(t) \). Similar to Section 3.1, a natural approach would use these same equations, only now estimating \( F_p(t) \) with the Kaplan-Meier estimate. In practice, however, this approach may not be feasible because \( \hat{F}_p(t) \) is not well defined for \( t > \tau \). The integral in the second estimating equation of (5) cannot be computed for \( t > \tau \).

One way to fix this problem would be to assume some distributional form for the tail of
$F_p(\cdot)$. This approach would be similar to the parametric methods of Section 3.2, however, and in this section we want to leave $F_p(\cdot)$ unspecified. Another approach would change the form of $w(\cdot)$, making it constant after time $\tau$. For example, consider $w(\cdot)$ defined as follows:

$$w(t; \alpha, \beta) = \begin{cases} 
(1 + \exp(-\alpha - \beta t))^{-1} & \text{for } t \leq \tau \\
(1 + \exp(-\alpha - \beta \tau))^{-1} & \text{for } t > \tau.
\end{cases} \quad (6)$$

Another choice for $w(\cdot)$ could be

$$w(t; \alpha, \beta) = \left(1 + \exp(-\alpha - \beta I\{t > t_0\})\right)^{-1} \quad (7)$$

for some $t_0 \leq \tau$. Both (6) and (7) define $w(\cdot)$ with the expit function, but do so in a manner such that $w(\cdot)$ is constant for $t > \tau$, allowing one to write:

$$\int_0^\infty w(s; \alpha, \beta) d\hat{F}_p(s) = \int_0^\tau w(s; \alpha, \beta) d\hat{F}_p(s) + w(\tau; \alpha, \beta) \left(1 - \hat{F}_p(\tau)\right).$$

Of course, these choices of $w(\cdot)$ have implications with regards to interpretation. Under (6) the interpretation of $\beta$ is technically the following: Given infection in the placebo arm, the odds of infection if randomized to the vaccine arm for $T = t_1$ versus $T = t_2$ are $\exp\{\beta [\min(t_1, \tau) - \min(t_2, \tau)]\}$. This more complex interpretation might appear troublesome. However, over the range of $t$ where there are data, $\beta$ has a standard log odds ratio interpretation. Under (7), $\beta$ has a standard odds ratio interpretation, except now we have dichotomized $w(\cdot)$, assigning a particular probability of being in the $ai$ stratum for $t > t_0$ and another for $t \leq t_0$.

With $w(\cdot)$ modeled by either (6) or (7), an extension of GBH using the Kaplan-Meier estimates for $F_p(t)$ is the semiparametric MLE. In the appendix we show that under these
same assumptions and $0 < VE < 1$, $p_0 > 0$, and for a properly specified well-behaved $w(\cdot)$ (i.e., constant for $t > \tau$, twice differentiable, and bounded), $\hat{F}_p(t)$ is consistent and asymptotically normal for $t \in (0, \tau]$. Therefore, $\hat{SCE}(t)$ is also consistent and asymptotically normal.

Because $\hat{SCE}(t)$ is asymptotically normal, pointwise Wald-based confidence intervals based on the bootstrap will be valid for large sample sizes. It is also possible to obtain an analytic form for the asymptotic variance of $\hat{SCE}(t)$. This variance estimate relies on being able to approximate $\hat{F}_p(t)$ with a sum of i.i.d. random variables. Such an approximation can be obtained from Stute (1995). Using this result, one can augment (5) by including additional estimating equations:

$$\{0, \cdots, 0\}^T = \left\{ \sum_{i=1}^{N} (1 - Z_i)S_i (V_{1i} - F_p(t_1)), \cdots, \sum_{i=1}^{N} (1 - Z_i)S_i (V_{ki} - F_p(t_k)) \right\}^T,$$

where $k$ is the number of distinct failure times in the placebo arm, $t_j$ is the $j$th ordered failure time, and for a specific $j$, $V_{ji}$ are i.i.d. random variables for $i = 1, \cdots, N$. One can then estimate the variance of parameter estimates using a sandwich estimator type approach, and from there one can estimate the variance of $\hat{SCE}(t)$ using the delta method. Details are given in the appendix.

4. SIMULATIONS

To evaluate the small-sample performance of our estimators of $SCE(t)$ we conducted a $2 \times 4$ factorial simulation experiment, corresponding to generating data under $VE \equiv P(S(1) = 0|S(0) = 1) \approx 0.3$ or 0.6, and $\beta = 0.1, 0.2, 1$, or $\infty$. Each simulation generated
1000 vectors $W$ according to the following steps:

**Step 1.** The first 500 vectors were set at $Z = 0$, the second 500 were set at $Z = 1$.

**Step 2.** $S(0)$ was drawn from a Bernoulli($p_0$) distribution with $p_0 = 0.25$ (this choice yields an expected number of infections in the placebo arm of 125, which is typical for a Phase III vaccine trial).

**Step 3.** $T(0)$ was generated for all realizations with $S(0) = 1$ according to the distribution $F_p(t; \eta)$ with $F_p(\cdot)$ a Weibull distribution and $\eta = (\text{shape} = 0.5, \text{scale} = 25)$. This distribution was chosen to reflect the distribution of the time from infection diagnosis to initiation of antiretroviral therapy (ART) in the VaxGen trial, in which approximately 50% of infected participants started ART by 24 months post-infection diagnosis.

**Step 4.** Given $T(0)$, for each realization with $Z = 1$ and $S(0) = 1$, $S(1)$ was drawn from a Bernoulli($w(T(0); \alpha, \beta)$) distribution. For $\beta = 0.1, 0.2$, and 1, $w(t; \alpha, \beta)$ was defined as in (6) with $\tau = 24$ months. To ensure that $VE \approx 0.3$, $\alpha$ was set at $-0.2, -0.9$, or $-3.6$, when $\beta$ was set at 0.1, 0.2, or 1 respectively; and to ensure that $VE \approx 0.6$, $\alpha$ was set at $-1.8, -3.4$, or $-20$, when $\beta = 0.1, 0.2$, or 1. For $\beta = \infty$, $w(t) = I_{\{t \geq q\}}$ as discussed in Section 2, where $q_{0.3} = 3.18$ and $q_{0.6} = 21.0$.

**Step 5.** For realizations with $Z = 1$ and $S(1) = 1$, $T(1)$ was set equal to $T(0)$.

**Step 6.** For the realizations with $S = 1$, $C_1$ was generated from a Weibull distribution with shape and scale parameters 3 and 35. Then $C$ was set as $\min(\tau, C_1)$.

**Step 7.** For all realizations with $S = 1$, $Y$ was chosen as $\min(C, T)$ and $\delta = I_{\{Y = T\}}$. 
It is easily verified that these steps result in simulating data under SUTVA, (1), A.1, A.2, A.3, and model (6), with $SCE(t) = 0$ for all $t$.

For each simulated dataset, we computed $\hat{F}_{ai}(t)$ and $\hat{SCE}(t)$ for $t = 24$ months, assuming the true model for $w(\cdot)$ and the true value for $\beta$. Wald-based 95% confidence intervals were constructed by estimating the standard error of estimates using both the bootstrap (with 200 bootstrap replications) and asymptotic variance estimates.

Table 1 reports the performance of $\hat{SCE}(t)$ based on 1000 simulation iterations. Since $\hat{SCE}(t)$ is the difference between $\hat{F}_{ai}(t)$ and $\hat{F}_{v}(t)$, it is also useful to examine the performance of the estimates of $F_{ai}(t)$. Table 2 does this using the same simulations and analyses reported in Table 1. In addition to presenting the coverage of the untransformed 95% Wald-confidence intervals for $F_{ai}(t)$, Table 2 presents confidence intervals by transforming the symmetric confidence limits for $\log[-\log\{F_{ai}(t)\}]$.

In most cases, bias is minimal and coverage is good using either the bootstrap or the asymptotic variance estimate. The only exceptions are when $VE \approx 0.6$ and $\beta$ is large. The poor coverage probabilities here are due to a boundary issue. Consider first the simulations with $\beta = \infty$. As discussed in Section 3.1, as $N \to \infty$, under the usual assumptions and if $VE < F_{p}(t)$, then $\hat{F}_{ai,L}(t)$ (which is equivalent to $\hat{F}_{ai}(t)$ with $\beta = \infty$) will be asymptotically normal and Wald-based confidence intervals will cover at their nominal level. However, in these simulations $F_{p}(t) = 0.625$ for $t = 24$ months, which is very close to $VE = 0.600$. Therefore, due to stochastic variation and our relatively small sample size, $\hat{VE}$ is often greater than $\hat{F}_{p}(t)$ resulting in $\hat{F}_{ai}(t) = 0$ in nearly half of the simulations. Consequently,
the distribution of $\hat{F}_{p}^{ai}(t)$ is far from normal; therefore these confidence intervals that assume normality have poor coverage. (Note that in this particular setting, Wald-based confidence intervals extended outside the $[0,1]$ range.) Interestingly, the 2.5- and 97.5- bootstrapped percentile confidence intervals for $F_{p}^{ai}(t)$ and $SCE(t)$ have coverage probabilities of 0.938 and 0.964, respectively, for the simulations with $VE \approx 0.6, \beta = \infty$.

The same logic explains why coverage and bias were also poor for the simulations with $VE \approx 0.6, \beta = 1$. Under these settings with time measured in months, a value of $\beta = 1$ is quite large: for example, the odds of being infected in the vaccine arm given infection in the placebo arm from $t = 12$ and $t = 24$ (a difference of one year) multiplicatively increase $exp(12) \approx 163,000$! Analyses based on the assumption that $\beta = 1$ are very similar to analyses assuming $\beta = \infty$. Again, the distribution of $\hat{F}_{p}^{ai}(t)$ is far from normal. Figure 1 shows a histogram of $\hat{F}_{p}^{ai}(t = 24)$ for the simulations with $VE \approx 0.6, \beta = 1$, as well as a similar histogram with $VE \approx 0.3, \beta = 1$ (where the method worked well) for purpose of comparison. Figure 1 also shows the true values of $F_{p}(t)$ and $F_{p}^{ai}(t)$ under these simulation settings. Notice how close $F_{p}^{ai}(t)$ is to 0 at $t = 24$ months.

5. EXAMPLE

We illustrate our methods using data from the VaxGen vaccine trial. This was a randomized, double-blind, placebo-controlled Phase III trial of AIDSVAX B/B conducted between 1998 and 2003. This study recruited 5,403 HIV negative, at risk individuals from 61 sites spanning large cities of North America and the Netherlands. The ratio of vaccine to placebo assignment was 2:1. Overall, the vaccine was not found to protect against HIV infection
(\hat{VE} = 0.048), although interaction tests suggested that the vaccine might partially prevent infection for non-whites. Among non-whites vaccine efficacy was estimated as 0.469. Detailed study results are found in Flynn et al. (2005). Here we compare the time from infection diagnosis to the initiation of antiretroviral therapy (ART) between the vaccine and placebo arms among participants (overall and within the non-white subgroup) who would have been infected regardless of randomization assignment. Specifically, we perform sensitivity analyses to test the hypothesis \( H_0 : SCE(t) = 0 \) for \( t = 2 \) years. A vaccine effect to delay ART is beneficial to an individual because it delays exposure to drug toxicities, drug resistance, and the depletion of future therapy options.

A total of 368 subjects were infected during the trial, and of these, 347 enrolled into the post-infection study phase (225 in the vaccine arm). There was presumably little interaction between trial participants, so SUTVA seemed reasonable. As discussed in SGJR, A.1 was also thought to be plausible because a) this was a double-blinded placebo controlled trial so behavior should not be different relative to assignment to the other treatment, and b) the vaccine was designed in such a way that it could not become the virus, causing infection. Monotonicity is not consistently testable, but there was no evidence from the trial nor from pre-trial experiments suggesting its being violated. In addition, the censoring mechanism did not appear to be informative based on similar drop-out rates for participants with different levels of behavioral risk; hence there is no evidence that A.2 is violated.

Figure 2 shows analyses looking at the time from infection diagnosis to the initiation of ART. Figure 2A shows the Kaplan-Meier estimates for both the vaccine and placebo arms for
the probability of not yet starting ART. The plot also includes the estimates of the upper and lower bounds of $F_p^{ai}(t)$, described in Section 3.1. Figure 2B shows a semiparametric sensitivity analyses looking at $SCE(t)$ for $t = 2$ years. The plot contains both the estimate for $SCE(t)$ and 95% Wald confidence intervals (constructed using the asymptotic variance approximation and the bootstrap, with 500 bootstrap replications). In these analyses (and all other sensitivity analyses in this section), we modeled the probability of infection in the vaccine arm given infection in the placebo arm, $w(t; \alpha, \beta)$, with (6). Before performing these analyses, we elicited a plausible range for the sensitivity parameter $\beta$ from a subject matter expert, Dr. Marc Gurwith of VaxGen. His best “guess” for a range for $exp(\beta)$ was 0.70 to 1.1, corresponding to $\beta$ from -0.36 to 0.1. (Note that $\beta$ is defined here in terms of years.) Figure 2B shows the estimates of $SCE(t)$ over a much larger range, for $\beta$ from -3 to 3. The open circles (and ‘+’ and ‘×’ signs) in the plots represent the sharp bounds of $SCE(t)$ (and 95% Wald and percentile confidence intervals, respectively, for these bounds based on 500 bootstrap replications), corresponding to an analysis with $\beta = \pm \infty$. (Confidence intervals based on the analytic variance were similar.) Regardless of the range, Figure 2 indicates that there is little evidence that the vaccine is having a causal effect on initiation of ART.

It may be of more interest to look at the effect of vaccination on the time-to-event outcomes for the non-white subgroup, where the vaccine appeared to partially protect against HIV infection. These analyses are shown in the lower plots of Figure 2. Because the estimate for $VE$ is much larger in the non-white cohort, the bounds for $F_p^{ai}(t)$ are farther apart, as seen in Figure 2C. This is also reflected in plot D, as the estimates for $SCE(t)$ cover a wider
range. Of course, the smaller sample size \((N = 914\) non-whites, of which 59 became infected) also inflates the length of the confidence intervals. It should be noted that 95% Wald-based confidence intervals of \(SCE(t)\) at \(\beta = \infty\) (where \(\hat{F}_{ai,L}(t) = 1\)) using the analytic variance expression were much wider than the bootstrap confidence intervals.

Notice that for these analyses in the non-white cohort, if \(\beta > 0\) then \(H_0 : SCE(t) = 0\) for \(t = 2\) years is rejected at the 0.05-level. This means that given infection in the placebo arm, if the odds of infection if randomized to vaccine are greater for someone who has a longer time to the initiation of ART, then there is evidence that the vaccine is causing non-white participants to have a higher probability of starting ART by 2 years post-infection diagnosis. This would imply that among non-whites the vaccine is having a detrimental effect, causing more rapid post-infection progression. Interestingly, this value of \(\beta\) is just inside Dr. Gurwith’s plausible range for non-whites, -0.92 to 0.18. Therefore, both the null and the alternative are favored in this range. However, it is worth noting that nowhere in this range is the point estimate positive; hence, the analysis provides no support for the vaccine being effective in non-whites.

6. DISCUSSION

As candidate vaccines continue to be developed and enter clinical trials, there is particular interest in looking at the effect of vaccination on post-infection outcomes. In this paper, we have proposed sensitivity analysis methods for evaluating the causal effect of vaccination on outcomes defined as the time from infection diagnosis to some post-infection event. We developed nonparametric, parametric, and semiparametric estimation methods, and applied
them to investigate the causal effect of VaxGen’s AIDSVAX B/B vaccine on the time from infection diagnosis to initiation of ART.

From our VaxGen analysis in the non-white cohort, different conclusions were drawn over the chosen range. The fact that our analyses produce more than one answer may not be attractive. We believe, however, that it is an honest way to present the data and allows scientists to look at results and draw their own conclusions. Based on our experience with Dr. Gurwith, eliciting a range for sensitivity parameters is feasible, although admittedly, different subject matter experts may choose different ranges for $\beta$. Nevertheless, from our communication with other HIV vaccine experts, those with worse post-infection outcomes in the placebo arm have generally been thought to be those most likely infected if randomized to vaccine (in our example, corresponding with $\beta \leq 0$), because either they are thought to have a weaker immune system or to have been exposed to a more virulent virus (Shepherd, Gilbert, Mehrotra, in press).

The approach discussed here ignores the time of infection diagnosis and any variation in the amount of time non-infected participants were followed. An alternative approach would be to use principal strata of time until infection diagnosis, which could be a better approach if there are reasons to believe that the causal effect of the vaccine on the post-infection outcome may differ based on the timing of infection. This could happen, for example, if those infected early in the trial are different immunologically than those infected later. Suppose $U$ is the discrete failure time from randomization to infection diagnosis, taking values \{$u_1, \ldots, u_K, u_{K+}$\}, where $K+$ indicates not being diagnosed with infection
by time $u_K$. Principal strata could be defined as the sets $PS_{jk} = \{U(0) = u_j, U(1) = u_k\}$, for $j, k = 1, \cdots, K, K+; one set of causal estimands of potential interest would be $SCE(t|PS_{jj}) \equiv P(T_i(0) \leq t|PS_{jj}) - P(T_i(1) \leq t|PS_{jj}), j = 1, \cdots, K$. (Note that in the absence of pre-infection censoring, the principal stratum focused on in this paper, $\{S(0) = S(1) = 1\}$, is the union of the time-dependent principal strata, $\cup_{j,k=1,\cdots,K} PS_{jk}$.) Due to sparseness of data within some principal strata, one might need to group times. For example, if there were $K = 6$ discrete failure times, one could focus on the principal strata defined by $\{U(0) \leq u_2, U(1) \leq u_2\}, \{u_2 < U(0) \leq u_4, u_2 < U(1) \leq u_4\},$ and $\{u_4 < U(0) \leq u_6, u_4 < U(1) \leq u_6\}$. Alternatively, one could focus on the principal strata defined by $PS(u) \equiv \{U(0) \leq u, U(1) \leq u\}$. For fixed $u$ this latter principal stratum is equivalent to the always infected principal stratum that would arise by redefining $S = 1$ as the indicator of infection by time $u$ post-randomization. Estimation of $SCE(t|PS(u))$ when there is censoring before time $u$ warrants further study.

It should be noted that the asymptotic normality of our estimators relies on the assumption that $VE > 0$. For $VE$ near the boundary $VE = 0$, Wald-based confidence intervals may have poor coverage (see Jemiai and Rotnitzky (submitted, 2005) for further discussion). In the VaxGen trial, using the entire cohort the estimate for $VE$ was 0.048. We performed an additional simulation investigating the performance of Wald-based 95% confidence intervals of $SCE(t = 24$ months$)$ under the same settings as described in Section 5 only generating data under the assumption that $VE = 0.05$ and $\beta = 0.1$. Coverage using the bootstrap and analytic variance estimates was good, 0.941 and 0.938, respectively. Of course, in the Vax-
Gen trial we do not know the true $VE$ so it may be misleading to read too much into these simulations. This is not an issue for the analysis of the non-white cohort where $\sqrt{VE} = 0.469$. However, the asymptotic normality of $\hat{F}_p^{n.i.}(t)$ for $t = 2$ years could be questioned in the non-white cohort since $\sqrt{VE} > \hat{F}_p(t)$.

Although these methods were designed specifically for HIV vaccine trials, they may apply in other situations where causal comparisons conditioning on post-randomization variables are of interest.

ACKNOWLEDGEMENTS

We would like to thank VaxGen for allowing us to use their data, and particularly Marc Gurwith for giving us ranges for the sensitivity parameter. This work was supported in part by NIH grant 2 RO1 AI054165-04.

REFERENCES


Diseases, 191, 647–649.


Rosenbaum, P.R. (1984), “The consequences of adjustment for a concomitant variable that has been affected by the treatment,” *Journal of the Royal Statistical Society, Series A*, 147, 656–666.


APPENDIX: TECHNICAL DETAILS

A.1 Asymptotic Normality Semiparametric Estimator

Suppose data are collected over a finite interval [0, \( \tau \)] with \( \tau \) fixed as \( N \to \infty \). Assume SUTVA, (1), A.1, A.2, A.3, \( 0 < VE < 1, p_0 > 0 \), and let \( t \in [0, \tau] \). Furthermore, assume \( w(t; \alpha, \beta) \in (0, 1) \) for all \( t > 0, \alpha, \) and \( \beta; \) is constant for \( t > \tau \); and is twice continuously
differentiable with respect to $\alpha$, with a bounded second derivative. Then $\hat{F}_p(t)$ is consistent and asymptotically normal.

**Proof.** Let $(\hat{p}_0, \hat{\alpha})$ be the solutions to the estimating equations given by (5). From the first equation, $\hat{p}_0 = n_p/N_p$, and the second equation can be written as

$$ U_N(\alpha) \equiv \int_0^\infty w(t; \alpha, \beta) d\hat{F}_p(t) - \left(1 - \hat{V}E\right) = 0. $$

Using the empirical distribution of $O_t$, the process $\hat{F}_p(\cdot)$ in $D[0, \tau]$ and $\hat{V}E$ are jointly asymptotically normal, where $D[0, \tau]$ is the space of cadlag functions (i.e., right continuous whose limits from the left exist) on $[0, \tau]$. The map $\left(\hat{F}_p(\cdot), \hat{V}E\right) \mapsto U_N(\alpha)$ is Hadamard differentiable (follows from Problem 7, Chapter 20, van der Vaart (1998)); therefore, by the functional delta method, $U_N(\alpha)$ is asymptotically normal. From Taylor’s expansion write

$$ \sqrt{N}(\hat{\alpha} - \alpha) = \frac{-\sqrt{N}U_N(\alpha)}{U_N'(\alpha) + \frac{1}{2}U_N''(\alpha^*)(\hat{\alpha} - \alpha)^2}, $$

where $\alpha^*$ is some value between $\alpha$ and $\hat{\alpha}$, and $U_N''(\alpha^*) = \int_0^\infty w''(t; \alpha^*, \beta) d\hat{F}_p(t)$. Notice that $U_N''(\alpha^*)$ is bounded because $|w''(t; \alpha, \beta)|$ is bounded. Therefore, because $U_N'(\alpha) \to^P U'(\alpha) \equiv \int_0^\infty w'(t; \alpha, \beta) dF_p(t)$ by the continuous mapping theorem, $\hat{\alpha} \to^P \alpha$ by Lemma 5.10 of van der Vaart (1998), $U_N''(\alpha^*)$ is bounded, and $U_N(\alpha)$ is asymptotically normal; $\hat{\alpha}$ is asymptotically normal. Finally, consider the map:

$$ \left(\hat{\alpha}, \hat{F}_p(\cdot)\right) \mapsto \frac{\int_0^t w(s; \hat{\alpha}, \beta) d\hat{F}_p(s)}{\int_0^\infty w(t; \hat{\alpha}, \beta) dF_p(t)} \equiv \hat{F}_p^{ai}(t). $$

That the process $\hat{F}_p^{ai}(\cdot)$ in $D[0, \tau]$ is asymptotically normal follows from Hadamard differentiability of this map, the chain rule, and the functional delta method. \qed
A.2 Asymptotic Variance of Semiparametric Estimator

From Stute (1995), \( \int \phi(t)d\hat{F}_p(t) \) can be written as a sum of i.i.d. terms plus a remainder term, \( R_{np} \), where \( |R_{np}| = o(n_p^{-1/2}) \), and \( \phi(t) \) is a well-behaved function of \( t \). Define \( k \) as the number of distinct failure times, and let \( t_1, \ldots, t_k \) represent the distinct ordered failure times. One could think of there being \( k+2 \) parameters to estimate, \((p_0, \alpha, F_p(t_1), F_p(t_2), \ldots, F_p(t_k)) \equiv \theta \), adding \( k \) estimating equations to (5):

\[
\psi_i(\theta) = \begin{cases} 
(1 - Z_i)(S_i - p_0) \\
Z_i (S_i - p_0 \int_0^\infty w(s; \alpha, \beta)dF_p(s)) \\
(1 - Z_i)S_i (V_{i1} - F_p(t_1)) \\
\vdots \\
(1 - Z_i)S_i (V_{ik} - F_p(t_k)) 
\end{cases}
\]

with \( V_{ji} \) for \( j = 1, \ldots, k \) and \( i = 1, \ldots, N \) defined as

\[
V_{ji} = \phi_j(Y_i) \gamma_0(Y_i) \delta_i + \gamma_{j1}(Y_i)(1 - \delta_i) - \gamma_{j2}(Y_i), \text{ where}
\]
\[
\phi_j(Y_i) = I_{\{Y_i \leq t_j\}},
\]
\[
\gamma_0(Y_i) = \exp\left(\int_{-\infty}^{Y_i} \frac{H^0(dz)}{1 - H(z)}\right),
\]
\[
\gamma_{j1}(Y_i) = \frac{1}{1 - H(Y_i)} \int I_{\{Y_i < \omega\}} \phi_j(\omega) \gamma_0(\omega) H^1(d\omega),
\]
\[
\gamma_{j2}(Y_i) = \int \int \frac{I_{(\nu < Y_i, \nu < \omega)} \phi_j(\omega) \gamma_0(\omega)}{(1 - H(\nu))^2} H^0(d\nu) H^1(d\omega),
\]
\[
H^0(y) = \mathbb{P}_N(Y \leq y, \delta = 0) = \frac{\sum (1 - z_i) s_i (1 - \delta_i) I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i},
\]
\[
H^1(y) = \mathbb{P}_N(Y \leq y, \delta = 1) = \frac{\sum (1 - z_i) s_i \delta_i I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i},
\]
\[
H(y) = \mathbb{P}_N(Y \leq y) = \frac{\sum (1 - z_i) s_i I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i}.
\]

To be clear and to simplify further notation:
\[
\int_0^{\infty} w(s; \alpha, \beta) dF_p(s) = \sum_{j=1}^{k+1} w_j(\alpha) (F_p(t_j) - F_p(t_{j-1})),
\]
where \( w_j(\alpha) = w(t_j; \alpha, \beta), w_{k+1}(\alpha) = w(t; \alpha, \beta), F_p(t_0) = 0, \) and \( F_p(t_{k+1}) = 1. \) From Appendix A.1, \( \sqrt{N} \left( \hat{\theta} - \theta \right) \to_d \mathcal{N}(0, \Psi), \) where
\[
\Psi = E \left[ \frac{\partial}{\partial \theta} \psi(\theta) \right]^{-1} E \left[ \psi(\theta) \psi(\theta)^T \right] E \left[ \left( \frac{\partial}{\partial \theta} \psi(\theta) \right)^T \right]^{-1}.
\]

Define
\[
g(\theta) = \frac{\sum_{j=1}^{l} w_j(\alpha) (F_p(t_j) - F_p(t_{j-1}))}{\sum_{j=1}^{k+1} w_j(\alpha) (F_p(t_j) - F_p(t_{j-1}))},
\]
where \( t_l = \sup(t_j) \) such that \( t_j < t. \) Therefore \( g(\hat{\theta}) = \hat{F}_p^{ni}(t). \) By the delta method,
\[
\sqrt{N} \left( g(\hat{\theta}) - g(\theta) \right) \to_d \mathcal{N}(0, g'(\theta) \Psi g'(\theta)^T).
\]

One may estimate \( g'(\theta) \) with \( g'(\hat{\theta}) \) and \( \Psi \) with
\[
\hat{\Psi} = \left[ \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial \theta} \psi_i(\hat{\theta}) \right]^{-1} \frac{1}{N} \sum_{i=1}^{N} \left[ \psi_i(\hat{\theta}) \psi_i(\hat{\theta})^T \right] \left[ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{\partial}{\partial \theta} \psi_i(\hat{\theta}) \right)^T \right]^{-1}.
\]
Table 1: Bias of estimates and coverage probability of Wald-based 95\% confidence intervals for $SCE(t)$ with $t = 24$ months.

<table>
<thead>
<tr>
<th>VE</th>
<th>$\beta$</th>
<th>Bias Mean</th>
<th>Bias Median</th>
<th>Coverage Probability</th>
<th>Bootstrap</th>
<th>Analytic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sim 0.3$</td>
<td>0.1</td>
<td>-0.002</td>
<td>0.000</td>
<td>0.946</td>
<td>0.948</td>
<td></td>
</tr>
<tr>
<td>$\sim 0.2$</td>
<td>-0.007</td>
<td>-0.007</td>
<td>0.943</td>
<td>0.949</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim 1$</td>
<td>-0.006</td>
<td>-0.003</td>
<td>0.943</td>
<td>0.946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim \infty$</td>
<td>-0.007</td>
<td>0.005</td>
<td>0.948</td>
<td>0.953</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim 0.6$</td>
<td>0.1</td>
<td>0.003</td>
<td>0.006</td>
<td>0.939</td>
<td>0.940</td>
<td></td>
</tr>
<tr>
<td>$\sim 0.2$</td>
<td>0.013</td>
<td>0.015</td>
<td>0.945</td>
<td>0.945</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim 1$</td>
<td>0.042</td>
<td>0.020</td>
<td>0.933</td>
<td>0.912</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim \infty$</td>
<td>0.035</td>
<td>0.000</td>
<td>0.935</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Analytic coverage probabilities when $VE \approx 0.6, \beta = \infty$ are not given because in many simulations, there were no failures in the vaccine arm.
Table 2: Bias of estimates and coverage probability of Wald-based 95% confidence intervals for $F_p^{ai}(t)$ with $t = 24$ months.

<table>
<thead>
<tr>
<th>VE</th>
<th>$\beta$</th>
<th>Bias</th>
<th>Coverage Probability</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Standard</td>
<td>Log-log transformation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Bootstrap</td>
<td>Analytic</td>
<td>Bootstrap</td>
<td>Analytic</td>
</tr>
<tr>
<td>$\sim 0.3$</td>
<td>0.1</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.939</td>
<td>0.941</td>
<td>0.946</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>-0.006</td>
<td>-0.002</td>
<td>0.932</td>
<td>0.939</td>
<td>0.938</td>
<td>0.949</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.001</td>
<td>-0.005</td>
<td>0.940</td>
<td>0.953</td>
<td>0.959</td>
<td>0.973</td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>-0.007</td>
<td>0.002</td>
<td>0.934</td>
<td>0.947</td>
<td>0.957</td>
<td>0.967</td>
</tr>
<tr>
<td>$\sim 0.6$</td>
<td>0.1</td>
<td>0.001</td>
<td>0.003</td>
<td>0.939</td>
<td>0.942</td>
<td>0.947</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.015</td>
<td>0.011</td>
<td>0.929</td>
<td>0.927</td>
<td>0.938</td>
<td>0.937</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.045</td>
<td>0.010</td>
<td>0.903</td>
<td>0.819</td>
<td>0.902</td>
<td>0.859</td>
</tr>
<tr>
<td></td>
<td>$\infty$</td>
<td>0.036</td>
<td>-0.016</td>
<td>0.889</td>
<td>0.948</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

NOTE: Analytic coverage probabilities when $VE \approx 0.6$, $\beta = \infty$ are not given because in many simulations, log[-log]-transformed confidence intervals could not be computed because the estimated value of $F_p^{ai}(t)$ was 0.
Figure 1: Histograms of $F_{ai}^p(t)$ for $t = 24$ months and plots of $F_p(t)$ and $F_{ai}^p(t)$ at different levels of $VE$ for $\beta = 1$. 
Figure 2: Sensitivity analyses of the causal effect of vaccination on the probability of not yet initiating antiretroviral therapy in the VaxGen trial (A and B) and in the trial’s non-white cohort (C and D). In plots B and D, $SCE(t = 2$ years), $\cdot$ $\cdot$ $\cdot$ is the 95% bootstrap Wald confidence interval (CI), $\cdot$ $\cdot$ $\cdot$ is the 95% CI based on the asymptotic variance, $\circ$ is $SCE(t)$ at $\beta = \pm \infty$, $+$ is 95% bootstrap Wald CI at $\beta = \pm \infty$, $\times$ is 95% bootstrap percentile CI at $\beta = \pm \infty$, and $[ ]$ is Dr. Gurwith’s plausible range for $\beta$. 